

LIST OF INVENTORS' NAMES AND ADDRESSES

Yuri HIRAIWA, Tokyo, JAPAN;
Hirofumi NAGASUKA, Tokyo, JAPAN;
Masahiro KIYOI, Tokyo, JAPAN;
Kazuhisa MIYATA, Tokyo, Japan;
Taketoshi SAKURABA, Tokyo, Japan.

ASA-1154
W1165-01EF

Title of the Invention

DISK SUBSYSTEM, COMPUTER SYSTEM, STORAGE MANAGING METHOD AND
PROGRAM

Inventors

Yuri HIRAIWA
Hirofumi NAGASUKA
Masahiro KIYOI
Kazuhisa MIYATA
Taketoshi SAKURABA

- 1 -

DISK SUBSYSTEM, COMPUTER SYSTEM, STORAGE
MANAGING METHOD AND PROGRAM

BACKGROUND OF THE INVENTION

The present invention relates to storage managing methods and more particularly, to a storage managing method which imposes less loads on a host side
5 computer and a network when management information is collected.

To meet the increase in amounts of data in mission critical systems or key systems of public organizations/enterprises assuming the responsibility
10 of the recent infrastructure, the storage system has coped with the increase in capacity of storage by using a technique for control of disk array. JP-A-2001-256003 discloses such a technique using a disk array.

Also, there has arisen a tendency to
15 establish the key system as above by using a SAN (Storage Area Network) representing a network dedicated to a storage, thereby coping with increased capacity and complexity of the storage. In this type of system, it is desirable that flexible management and expansion
20 of the storage be implemented by connecting a server and the storage by the SAN so as to manage the storage independently of the server. JP-A-2002-7304 discloses a technique concerning a method for managing such a system as using the SAN. By virtue of this known

technique, such a storage device as RAID (Redundant Arrays of Inexpensive Disks) or MT (Magnetic Tape) can be worked unitarily or as a single element.

Further, JP-A-2002-175203 discloses a work
5 method directed to load distribution in such a SAN environment that a system has a plurality of servers and a plurality of disk devices are allocated to each of the plurality of servers. This known technique aims at distributing loads or load balancing in accessing
10 the disks and supervising resources for this purpose.

SUMMARY OF THE INVENTION

In the JP-A-2002-175203, when information is collected from a server processing business applications to a management server taking in charge of
15 resource management, the server processing business applications is provided with a program for performing resource supervisory and information about loads imposed on disk volumes is periodically collected from the program. Accordingly, the program for periodical
20 load information collection uses a CPU to impose a load on the CPU for processing of business applications.

As information necessary for unitary work in the SAN environment, information about availability (unused space) of disk, for instance, can be numerated
25 in addition to the access information described in JP-A-2002-175203. In case a method is employed which provides the server with a program for performing

resource supervisory with a view to acquiring the information as above, an instruction for acquisition of the information such as unused space information must be issued from the server side to the disks.

- 5 Accordingly, there is a possibility that an I/O instruction for processing business applications is delayed.

The present invention has been made to solve the above problems and it is an object of the invention
10 to provide a computer system which can alleviate loads imposed on a host computer and a network when information for management of a storage of a disk subsystem is collected, thereby permitting efficient access to the storage.

- 15 According to one feature of the invention, a computer system comprises a disk subsystem provided with a CPU, a business service computer for consulting the disk subsystem and a management computer for managing the system, which are mutually connected by a
20 network, wherein an information collecting program (also called "agent") for collecting information about a disk in accordance with a request from the management computer and reporting the collected information runs on the disk subsystem provided with the CPU.

- 25 According to a feature of the computer system of the invention, the agent is operated on the disk subsystem, so that information about unused space of the disk can be collected without imposing a load on a

computer for processing business applications. In
collecting the disk information, the unused space
information can be transmitted to the management
computer without being routed through the business
5 service computer and as a result, the disk unused space
information can be collected without imposing a load on
a communication path between that computer and the
disk.

Other objects, features and advantages of the
10 invention will become apparent from the following
description of the embodiments of the invention taken
in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram showing the
15 construction of a computer system according to the
invention.

Fig. 2 is a diagram showing the structure of
volume management information 200 and the structure of
a volume.

20 Fig. 3 is a diagrammatic representation of
volume designation information 300.

Fig. 4 is a diagrammatic representation of
volume configuration information 400.

Fig. 5 is a diagrammatic representation of
25 volume management information extracted data 500.

Fig. 6 is a flowchart showing the flow of a
process of supervisory program 600.

DESCRIPTION OF THE EMBODIMENTS

The embodiments of the present invention will now be described by way of example with reference to the accompanying drawings.

5 The present invention is concerned with a storage managing method. Conceivably, various kinds of data are handled for storage management but in particular, the present embodiment will be described on the assumption that data of a disk volume needed by a
10 management program has the contents, three in total, of disk unused area capacity, total capacity and access frequency.

[Configuration of Computer System]

Referring first to Fig. 1, description is
15 made of a computer system according to the embodiment of the invention in configuration.

The present system as shown in Fig. 1 comprises a management computer 100, computers 800 (800-1, 800-2) of objects to be managed and a disk
20 subsystem 900 accessible from the computers. These components are connected to each other through a network. Attached to the disk subsystem 900 is a SVP (Service Processor) 1000 for performing management of disks inside the disk subsystem 900. The computers 800
25 are connected to the disk subsystem 900 through the medium of a SAN 1100 representing a network dedicated to storage. On the other hand, the management computer 100 is connected to the computers 800 and SVP 1000

through a LAN (Local Area Network) 1200.

An AP (Application Program) 803 (803-1, 803-2) for processing business applications by users is loaded on a main memory 802 (802-1, 802-2) of computer 5 800 (800-1, 800-2) and executed by means of a CPU 801 ((801-1, 801-2). In order to conduct business affairs, the AP 803 accesses the disk subsystem 900 to read and write data file by file, as necessary.

The disk subsystem 900 incorporates disks 904 10 (904-1, 904-2) controlled by means of a disk controller 903. The disk 904 may be a virtual disk (logical volume) and each disk may be an array comprised of two or more disks. In the present embodiment, I/O instructions for disks 904 are identified by disk ID's 15 and are allocated as logical volumes in unit of disk to each computer 800, thereby ensuring that collection/management of information can be carried out in unit of disk.

Volume management information 200 (200-1, 20 200-2) of a format depending on a file system of an allocated computer 800 is stored on the disk 904.

On a main memory 902 of disk subsystem 900, an agent 700 for performing resource supervisory is developed and executed by means of a CPU 901. The 25 agent 700 is a program for collecting management information of the volume installed on the disk subsystem and it includes volume designation information 300 for designating a volume which delivers

management information, volume configuration
information 400 describing a method for referencing the
volume and supervisory program 600. On the basis of
the two pieces of volume designation information 300
5 and volume configuration information 400, the
supervisory program 600 delivers the management
information to volume management information extracted
data (area) 500 located on a main memory 1002 of SVP
1000. Then, in response to a request from the
10 management computer 100, the SVP 1000 transmits the
volume management information extracted data 500
through the LAN 1200.

The agent 700 is normally stored in an
auxiliary memory 1003 the SVP 1000 has and is
15 developed, as necessary, from the SVP 1000 onto the
main memory 902 of disk subsystem 900 so as to be
executed.

Even after the development, the volume
designation information 300 can be updated from the SVP
20 1000 to meet a change in configuration of disk 904. A
decision as to whether or not the information
acquisition is to be carried out can also be changed.

On a main memory 102 of management computer
100, a management program 103 is developed and executed
25 by means of a CPU 101. The management program 103
reads the volume management information extracted data
500 of SVP 1000 to collect the information of the disk
subsystem 900.

[Data Structure for Storage Management]

Next, a data structure for storage management according to an embodiment of the invention will be described with reference to Figs. 2 to 5, of which Fig. 2 illustrates volume management information 200 and the structure of a volume, Fig. 3 diagrammatically depicts volume designation information 300, Fig. 4 diagrammatically depicts volume configuration information 400 and Fig. 5 diagrammatically depicts volume management information extracted data 500.

The volume management information 200 is data for managing a volume corresponding thereto and is located on a disk 904 as shown in Fig. 2.

The disk 904 has head position 209 indicative of a start position of the volume management information 200. Disk ID 201 is an area adapted to store an ID for identifying the volume. Unused (area) capacity 202 is an area storing an unused (area) capacity of the volume. Total capacity 203 is an area storing a total capacity of the volume. Pointer 204 to a head control block is an area storing a pointer to the head of a list of control blocks 205 (205-1, 205-2) for managing areas used by files.

The control blocks 205 form a table having a list structure for managing data blocks used by files in the volume and each control table includes file information 206, pointer 207 to next control block and pointer 208 to a data block. The file information 206

is an area adapted to store information for identifying a file using the data block. This information sometimes indicates an unused (or free) status. The pointer 207 to next control block is an area adapted to
5 store a pointer to the next control block 205. The pointer 208 to a data block is an area adapted to store a pointer to a data block in the volume.

The volume designation information 300 describes information for collecting management
10 information of the volume and as shown in Fig. 3, it includes disk ID 301, information collection necessity/non-necessity 302 and volume configuration ID 303.

The disk ID 301 is an area storing ID's for
15 identifying volumes. The information collection necessity/non-necessity 302 is an area storing designation as to whether information collection is to be carried out in respect of a particular volume and in accordance with a value in this area, the supervisory
20 program decides whether management information of the volume is to be collected.

In the present embodiment, it is assumed that the area is described with "1" in the case of execution of collection and with "0" in the case of inexecution
25 of collection. The volume configuration information 400 is an area adapted to store information for indicating information as to which ID of the volume configuration information 400 is used in order to read

information from the volume management information 200 of the volume.

The volume configuration information is obtained by extracting information necessary for volume
5 management and as shown in Fig. 4, it includes volume configuration ID 401, head position information 402, unused (area) capacity describing position 403 and total capacity describing position 404.

The volume configuration ID 401 is an area
10 storing an ID for identifying the kind of volume configuration and is linked with the volume configuration ID of volume designation information of Fig. 3.

The head position information 402 is
15 described with information as to which location in the volume the head position 209 representing the area storing the head position of volume management information 200 is located at. The unused (area) capacity describing position 403 is an area storing a
20 position of the unused (area) capacity 202 in the volume management information 200. The total capacity describing position 404 is an area storing a position of the total capacity 203 in the volume management information 200.

25 When in the notation of the present embodiment the unused (area) capacity describing position 403 and total capacity describing position 404 are so described as to be parenthesized, a difference

from the head of volume management information 200 obtained from the head position information 402 is indicated but when not parenthesized, they are so described as to indicate positions in the volume.

5 The volume management information extracted data 500 is information for the SVP 1000 to hold volume management information of a necessary volume and as shown in Fig. 5, and it includes disk ID 501, unused (area) capacity 502, total capacity 503 and access
10 frequency 504.

 The disk ID 501 is an area storing an ID for identifying a volume. The unused (area) capacity 502 is an area storing an unused (area) capacity of a disk ID 501 acquired by the supervisory program 600. The
15 total capacity 503 is an area storing a total capacity of the disk ID 501 acquired by the supervisory program 600. The access frequency 504 is an area storing a frequency of access to a volume of the disk ID 501 acquired by the supervisory program 600.

20 [Process of Supervisory Program]

 The flow of a process of supervisory program 600 will now be described by making reference to Fig. 6 showing a flowchart of the process flow.

 Firstly, the supervisory program acquires
25 from the volume designation information 300 a disk ID 301 for which the information collection necessity/non-necessity 302 is "1" (step 601).

 Then, it is decided whether the disk ID 301

acquired in the step 601 is not unused, that is,
whether a disk from which information is collected
exists (step 602).

When the disk from which information is
5 collected is present, a volume configuration ID 303
corresponding to the disk ID 301 is acquired (step 603)
but in the absence of the disk from which information
is collected, one round of the process of supervisory
program is determined to be ended and the next startup
10 moment is waited for (step 608).

When the disk from which information is
collected is present and the step 603 is executed,
information is acquired from the volume management
information 200 of a disk 904 indicated next by a disk
15 ID 301. Namely, a volume configuration ID 401
coincident with the volume configuration ID 303
acquired in the step 603 is found out of the volume
configuration information 400. Then, the head position
information 402, unused (area) capacity describing
20 position 403 and total capacity describing position 404
corresponding to the found volume configuration ID are
obtained and by using the two pieces of information,
values of unused (area) capacity 202 and total capacity
203 are acquired from the volume management information
25 200.

For example, when the disk ID 301 is "1", the
head position information 402 corresponding to a value
of "1" of the volume configuration ID 401 is "4 bytes".

In addition, the unused (area) capacity describing position 403 is "(4 bytes)" and the total capacity describing position 404 is "(8 bytes)". Accordingly, the unused (area) capacity 202 is acquired from a
5 position which is further shifted by 4 bytes from a position (start position of volume management information 200) which is indicated by a pointer described at 4 bytes spaced from the head of the disk 904 for which the disk ID 301 is "1" and similarly, a
10 total capacity 203 is acquired from a position which is 8 bytes shifted from the start position of volume management information 200 (step 604).

Further, access frequency information of the disk 904 indicated by the disk ID 301 is acquired (step
15 605).

Pieces of information acquired in the steps 604 and 605 are output to the volume management information extracted data 500 (step 606).

A disk ID to be processed next is acquired
20 from the volume designation information 300 (step 607). Then, for all disk ID's to be processed in the manner as above, the process covering the steps 602 through 607 is repeated.

[Other Embodiments]

25 In the foregoing embodiments, the invention is carried out in such a way that data needed by the management program is extracted from the volume management information stored in the disk.

Alternatively, start and end points of the volume management information may be designated in order that the whole of the volume management information can be output. In this case, the whole of the volume
5 management information 200 is stored in the volume management information extracted data 500.

Then, in such an instance as above, the volume management information is not output as a whole but a difference from the preceding output may be
10 output.

The management information is embodied as having three kinds of information of volume unused (area) capacity, total capacity and access frequency but in addition thereto, volume allocation information
15 by computer, RAID emulation type and frequency of erroneous I/O write in unit of disk may be conceivable.

It should be further understood by those skilled in the art that although the foregoing description has been made on embodiments of the
20 invention, the invention is not limited thereto and various changes and modifications may be made without departing from the spirit of the invention and the scope of the appended claims.